

# Sampling on the Sphere by Mutually Orthogonal Subspaces

Uri Grupel \*

## Abstract

The purpose of this paper is twofold. First, we provide an optimal  $\Omega(\sqrt{n})$  bits lower bound for any two-way protocol for the Vector in Subspace Communication Problem which is of bounded total rank. This result complements Raz's  $O(\sqrt{n})$  protocol, which has a simple variant of bounded total rank. Second, we present a plausible mathematical conjecture on a measure concentration phenomenon that implies an  $\Omega(\sqrt{n})$  lower bound for a general protocol. We prove the conjecture for the subclass of sets that depend only on  $O(\sqrt{n})$  directions.

## 1 Introduction

The Vector in Subspace Problem (VSP) is a communication problem where one party (Alice) receives a unit vector  $u \in S^{n-1}$ , and a second party (Bob) receives a subspace  $H \subseteq \mathbb{R}^n$  of dimension  $\lfloor n/2 \rfloor$  such that either  $u \in H$  or  $u \in H^\perp$ . The goal of Alice and Bob is to determine whether  $u \in H$  or not.

VSP was introduced by Kremer [6] and has been studied under both classical and quantum communication models. In the classical communication model, Alice and Bob exchange bits between them in order to determine whether  $u \in H$ . In the quantum communication model, Alice and Bob exchange qubits.

In this paper, the terms protocol and complexity refer to distributional complexity (see [12] [1]). That is, a protocol outputs the correct answer with probability at least  $2/3$ . The complexity is measured according to the number of bits or qubits that are exchanged in the worst case.

It is known that VSP can be solved in the quantum model with the exchange of  $O(\log n)$  qubits. In [8] Raz presented a classical protocol that solves VSP with the exchange of  $O(\sqrt{n})$  bits. In [4], Klartag and Regev proved that any classical protocol for VSP has a communication complexity of at least  $\Omega(n^{1/3})$  bits. Thus, VSP shows that quantum communication can be exponentially stronger than classical communication. In this paper we discuss the gap between the lower and upper bound for the classical model.

We focus on the class of protocols of *bounded total rank*. In a deterministic protocol, each decision by Alice on the value of the next bit to be sent to Bob is based on two factors: her knowledge of the communication received so far, and perhaps an additional measurement of the vector  $u$ . We define the rank of the decision to be the number of linear functionals of the vector  $u$  that Alice has to compute in order to carry out the measurement. For example, deciding the value of the next bit by using the indicator function of the set  $\{\langle x, v_1 \rangle \geq 0, \sin(\langle x, v_2 \rangle^2) \geq 1/2\}$ , where  $v_1, v_2 \in \mathbb{R}^n$  are determined by the communication received so far, is a decision of rank 2. In general, the rank of a decision is an integer between 0 and  $n$ . The total rank of a protocol is the sum of all ranks of decisions made by Alice in the worst case scenario. Note that we do not count decisions by Bob. The protocol Raz introduced can be slightly modified to be of total rank  $O(\sqrt{n})$  (for more details see Appendix A).

We prove that any protocol of VSP, for the classical model, of total rank at most  $O(\sqrt{n})$  has communication complexity of at least  $\Omega(\sqrt{n})$  bits. In light of the upper bound by Raz, this lower bound is sharp.

---

\*Weizmann Institute of Science. urigrupel@gmail.com

We also introduce a novel mathematical conjecture about concentration of measure in the high dimensional sphere. This conjecture implies that any classical protocol for VSP has a communication complexity of at least  $\Omega(\sqrt{n})$  bits.

The lower bound by Klartag and Regev is a result of a concentration theorem for sampling on the sphere by random subspaces. They proved that for any measurable subset  $A \subseteq S^{n-1}$  with  $\sigma_{n-1}(A) \geq Ce^{-cn^{1/3}}$ , where  $\sigma_{n-1}$  is the uniform probability measure on the sphere  $S^{n-1}$ , it holds that

$$\mathbb{P}_H \left( \left| \frac{\sigma_H(A \cap H)}{\sigma_{n-1}(A)} - 1 \right| \leq 0.1 \right) \geq 1 - e^{-c'n^{1/3}},$$

where  $C, c, c' > 0$  are universal constants. Here  $\sigma_H$  denotes the Haar probability measure on  $S^{n-1} \cap H$  and  $\mathbb{P}_H$  denotes the orthogonally invariant Haar probability measure over the Grassmanian manifold of subspaces  $H \subseteq \mathbb{R}^n$  of dimension  $\lfloor n/2 \rfloor$ .

The concentration inequality by Klartag and Regev is sharp. Taking  $A = \{x \in S^{n-1}; x_1 \geq T\}$ , where  $T \approx n^{-1/3}$  is chosen such that  $\sigma_{n-1}(A) = n^{-1/3}$  gives

$$\mathbb{P}_H \left( \left| \frac{\sigma_H(A \cap H)}{\sigma_{n-1}(A)} - 1 \right| \leq 0.1 \right) = 1 - e^{-c_n n^{1/3}},$$

where  $c_n$  has a finite limit  $c \in (0, \infty)$  as  $n \rightarrow \infty$ .

Our goal is to find a concentration inequality that applies to smaller sets, that is sets with measure of the order of magnitude of  $e^{-\sqrt{n}}$ . Our hope is that by considering both  $H$  and  $H^\perp$  simultaneously, a stronger concentration result can be achieved.

**Conjecture 1.1.** *Let  $A \subseteq S^{n-1}$  be a measurable subset with  $\sigma_{n-1}(A) \geq e^{-c\sqrt{n}}$ . Then*

$$\mathbb{P}_H \left( \sqrt{\sigma_H(A \cap H) \sigma_{H^\perp}(A \cap H^\perp)} \geq 0.9 \sigma_{n-1}(A) \right) \geq 1 - Ce^{-c'\sqrt{n}},$$

where  $C, c, c' > 0$  are universal constants.

Conjecture 1.1 was essentially suggested by Klartag and Regev [5], albeit with a weaker arithmetic average in place of the geometric one.

In §3 we prove a special case of the conjecture where the set  $A \subseteq S^{n-1}$  is of the form  $\{x \in S^{n-1}; (x_1, \dots, x_k) \in I\}$  for some Borel set  $I \subseteq B_k = \{x \in \mathbb{R}^k; |x| \leq 1\}$ , and  $k = O(\sqrt{n})$ . By considering the case  $k = 1$ , this result shows that the conjecture holds for the extremal case of the theorem by Klartag and Regev. This extremal case also shows that if the conjecture is true it is tight.

This special case of the conjecture follows from the following result:

**Theorem 1.2.** *Let  $k \leq \alpha_1 \sqrt{n}$ . Let  $f : S^{n-1} \rightarrow [0, \infty)$  be a measurable function such that  $\|f\|_\infty \leq e^{\alpha_2 \sqrt{n}}$ ,  $\|f\|_1 = 1$  and  $f$  depends only on  $x_1, \dots, x_k$ . Then,*

$$\mathbb{P}_H \left( \sqrt{\int_{S_H} f(x) d\sigma_H(x) \int_{S_{H^\perp}} f(x) d\sigma_{H^\perp}(x)} \geq 0.9 \right) \geq 1 - \alpha_3 e^{-\sqrt{n}},$$

where  $\alpha_1, \alpha_2, \alpha_3 > 0$  are universal constants.

In the proof we use various tools from Geometric Functional Analysis. We begin by reformulating the problem in terms of random matrices instead of random subspaces. We show that the event in Theorem 1.2 strongly depends on the singular values of a random projection. Using results from the theory of Wishart matrices we show that these singular values are concentrated around their expected values.

Next, we use the Cauchy-Schwarz inequality to define a smaller event than the one in the theorem. The use of the Cauchy-Schwarz inequality demonstrates how considering both  $H$  and  $H^\perp$  simultaneously can enhance the concentration results and is fundamental to our approach.

Finally, we use the concentration results, and asymptotic tools such as the Laplace method in order to show that with high probability this smaller event holds true. In this last step we present bounds for the universal constants in Theorem 1.2 which are, in principle, explicit.

In §2 we employ the rectangle method and show how Theorem 1.2 implies a sharp lower bound for classical protocols of total rank at most  $O(\sqrt{n})$ .

**Corollary 1.3.** *Let  $\mathcal{P}$  be a protocol for the Vector in Subspace Problem of total rank at most  $\alpha_1\sqrt{n}$ , with probability of error which is at most a constant smaller than  $\frac{1}{2}$ . Then the communication complexity of  $\mathcal{P}$  is  $\Omega(\sqrt{n})$  bits.*

Using the same methods, a positive resolution of Conjecture 1.1 would imply a sharp lower bound for a general classical protocol to VSP.

**Theorem 1.4.** *Let  $\mathcal{P}$  be a general protocol for the Vector in Subspace Problem with probability of error which is at most a constant smaller than  $\frac{1}{2}$ . If Conjecture 1.1 is true then the communication complexity of  $\mathcal{P}$  is  $\Omega(\sqrt{n})$  bits.*

*Acknowledgement.* This paper was written under the supervision of Bo'az Klartag whose guidance, support and patience were invaluable. In addition, I would like to thank Sasha Sodin for useful discussions and suggestions, and Oded Regev for his remarks on an early draft of this paper. Supported by the European Research Council (ERC).

## 2 Applications to VSP

In this section we prove that Theorem 1.2 implies a lower bound of  $O(\sqrt{n})$  for classical protocols of total rank at most  $O(\sqrt{n})$ . We also show that Conjecture 1.1 implies a lower bound of  $O(\sqrt{n})$  for any classical protocol. In light of the result of Raz, if the conjecture is true then this bound is sharp.

Theorem 1.2 implies a special case of the conjecture for sets that depend only on  $\alpha_1\sqrt{n}$  directions. For any such  $A \subseteq S^{n-1}$  with  $\sigma_{n-1}(A) \geq e^{-\alpha_2\sqrt{n}}$ , define  $f(x) = 1_A(x)/\sigma_{n-1}(A)$ . The function  $f$  depends only on  $\alpha_1\sqrt{n}$  directions, bounded by  $1/\sigma_{n-1}(A) \leq e^{\alpha_2\sqrt{n}}$  and has  $\|f\|_1 = 1$ . Hence, we may apply Theorem 1.2. We have,

$$\begin{aligned} 1 - \alpha_3 e^{-\sqrt{n}} &\leq \mathbb{P}_H \left( \sqrt{\int_{S_H} 1_A(x)/\sigma_{n-1}(A) d\sigma_H(x)} \sqrt{\int_{S_{H^\perp}} 1_A(x)/\sigma_{n-1}(A) d\sigma_{H^\perp}(x)} \geq 0.9 \right) \\ &= \mathbb{P}_H \left( \sqrt{\sigma_H(A \cap H) \sigma_{H^\perp}(A \cap H^\perp) / \sigma_{n-1}(A)} \geq 0.9 \right). \end{aligned}$$

In this section, we use this consequence of Theorem 1.2.

Our argument follows the rectangle method usually attributed to Babai, Frankl and Simon [1] and Razborov [9].

For simplicity, we assume that  $n$  is even. We denote by  $G_{n/2}$  the Grassmanian manifold of all subspaces of  $\mathbb{R}^n$  of dimension  $n/2$ , equipped with the  $O_n$  invariant measure  $\sigma_G$ . Let  $\mu_0$  be the uniform measure on  $S^{n-1} \times G_{n/2}$ . Denote

$$I_1 = \{(u, H) \in S^{n-1} \times G_{n/2}; u \in H\},$$

and

$$I_2 = \{(u, H) \in S^{n-1} \times G_{n/2}; u \in H^\perp\}.$$

Let  $\mu_i$  be the Haar invariant probability measure on  $I_i$  for  $i = 1, 2$ , with respect to the obvious  $O_n$  action. Such measure exists due to the transitive property of such action. For a rectangular set  $A_i \times B_i \subseteq I_i$  we have

$$\mu_1(A_1 \times B_1) = \int_{H \in B_1} \sigma_H(A_1 \cap H),$$

and

$$\mu_2(A_2 \times B_2) = \int_{H \in B_2} \sigma_{H^\perp}(A_2 \cap H^\perp).$$

By replacing  $\alpha_2$  in Theorem 1.2 with  $\min\{\alpha_2, 1\}$  we may assume that it is at most 1.

**Proposition 2.1.** *Let  $Q = A \times B \subseteq S^{n-1} \times G_{n/2}$  be such that  $\mu_0(A \times B) \geq Ce^{-\alpha_2\sqrt{n}}$ . Assume that the set  $A$  depends on  $\alpha_1\sqrt{n}$  directions. Then*

$$\sqrt{\mu_1(Q)\mu_2(Q)} \geq 0.8\mu_0(Q).$$

The constants  $\alpha_1, \alpha_2, C > 0$  are universal constants.

*Proof.* Define

$$E = \left\{ H \in B; \sqrt{\sigma_H(A \cap H)\sigma_{H^\perp}(A \cap H^\perp)} \leq 0.9\sigma(A) \right\}.$$

According to our assumption  $\sigma_{n-1}(A) \geq \mu_0(A \times B) \geq Ce^{-\alpha_2\sqrt{n}}$ . By the Theorem 1.2,  $\mathbb{P}(E) \leq \alpha_3 e^{-\sqrt{n}}$ . We choose  $C \geq 1$  big enough, such that

$$0.9\sigma_G(B \setminus E) \geq 0.8\sigma_G(B). \tag{1}$$

By the Cauchy-Schwarz inequality

$$\begin{aligned} \sqrt{\mu_1(Q)\mu_2(Q)} &= \sqrt{\left(\int_{H \in B} \sigma_H(A \cap H)\right) \left(\int_{H \in B} \sigma_{H^\perp}(A \cap H^\perp)\right)} \geq \int_{H \in B} \sqrt{\sigma_H(A \cap H)\sigma_{H^\perp}(A \cap H^\perp)} \\ &\geq \int_{H \in B \setminus E} \sqrt{\sigma_H(A \cap H)\sigma_{H^\perp}(A \cap H^\perp)} \geq 0.9 \int_{H \in B \setminus E} \sigma(A) = 0.9\sigma(A)\sigma_G(B \setminus E). \end{aligned}$$

Equation (1) gives us

$$\sqrt{\mu_1(Q)\mu_2(Q)} \geq 0.8\sigma_G(B)\sigma(A) = 0.8\mu_0(Q).$$

□

**Corollary 2.2.** *Let  $Q = A \times B \subseteq S^{n-1} \times G_{n/2}$ . Assume that the set  $A$  depends on  $\alpha_1\sqrt{n}$  directions. Then*

$$\sqrt{\mu_1(Q)\mu_2(Q)} \geq 0.8\mu_0(Q) - Ce^{-\alpha_2\sqrt{n}}.$$

Using the above propositions, we are ready to prove Corollary 1.3.

*Proof.* By repeated application of the protocol we may assume that the probability of error is less than  $\frac{1}{9}$ . By Yao's principle [11], we may assume that our protocol is a randomly chosen deterministic protocol. Let  $D$  be the number of bits exchange in the protocol. We have a partition of  $S^{n-1} \times G_{n/2}$  into  $2^D$  rectangles of the form  $Q = A \times B$ , each labeled as "In  $H$ " or "In  $H^\perp$ ". Since we assume the total rank is at most  $\alpha_1 \sqrt{n}$ , for every  $Q = A \times B$  in the partition, the set  $A$  is determined by at most  $\alpha_1 \sqrt{n}$  directions. Denote by  $\mathcal{Q}_+$  all the rectangles labeled "In  $H$ ", and by  $\mathcal{Q}_-$  all the rectangles labeled "In  $H^\perp$ ". According to our assumption  $\sum_{Q \in \mathcal{Q}_+} \mu_2(Q) \leq \frac{1}{9}$  and  $\sum_{Q \in \mathcal{Q}_-} \mu_1(Q) \leq \frac{1}{9}$ . By Corollary 2.2 and the Cauchy-Schwarz inequality, we have

$$\begin{aligned} \sum_{Q \in \mathcal{Q}_+} (0.8\mu_0(Q) - Ce^{-\alpha_2 \sqrt{n}}) &\leq \sum_{Q \in \mathcal{Q}_+} \sqrt{\mu_1(Q)\mu_2(Q)} \leq \sqrt{\left(\sum_{Q \in \mathcal{Q}_+} \mu_1(Q)\right) \left(\sum_{Q \in \mathcal{Q}_+} \mu_2(Q)\right)} \\ &\leq \sqrt{1 \cdot \frac{1}{9}} = \frac{1}{3}. \end{aligned}$$

Similarly we have,

$$\sum_{Q \in \mathcal{Q}_-} (0.8\mu_0(Q) - Ce^{-\alpha_2 \sqrt{n}}) \leq \frac{1}{3}.$$

Summing the above inequalities, we obtain

$$0.8 - 2^D Ce^{-\alpha_2 \sqrt{n}} \leq \frac{2}{3} \Rightarrow D \geq C' \sqrt{n}.$$

□

If Conjecture 1.1 is true, then Proposition 2.1 and Corollary 2.2 are true without the assumption that the set  $A$  depends on  $\alpha_1 \sqrt{n}$  directions. Hence, we may repeat the proof of Corollary 1.3 for a general classical protocol, and deduce Theorem 1.4.

### 3 Proofs

In this section we prove Theorem 1.2. This theorem is a special case of the conjecture for functions that depend only on  $O(\sqrt{n})$  directions. We assume that  $n$  is even and greater than some universal constant. In this section, when we say *uniform distribution*, we refer to the Haar probability distribution.

Throughout this section we shall use the letters  $c, \tilde{c}, C$  etc. to denote various universal constants, whose value may change from one line to the next. Additionally,  $\alpha_1, \alpha_2, \alpha_3, \rho > 0$  are universal constants whose value would be determined only at the end of the section. Specifically, in this section we will assume a few upper bounds for  $\alpha_1$  in terms of explicit positive universal constants, an upper bound for  $\alpha_2$  in terms of  $\alpha_1$ , and a lower bound for  $\alpha_3$  in terms of  $\alpha_2$ .

Let  $\psi : B_k \times S^{m-k-1} \rightarrow S^{m-1}$  be defined by  $\psi(x, y) = (x, \sqrt{1 - |x|^2}y)$ . The map  $\psi$  enables us to separate the dependence on the first  $k$  coordinates. The following change of variables formula is standard:

**Proposition 3.1.** *For any integrable  $f : S^{m-1} \rightarrow \mathbb{R}$  and any  $1 \leq k \leq m-1$  there exists  $C_{m,k} = (m-k) \text{Vol}(B_{m-k}) / (m \text{Vol}(B_m))$  such that*

$$\int_{S^{m-1}} f(x) d\sigma_{m-1}(x) = C_{m,k} \int_{B_k} (1 - |x|^2)^{(m-k-2)/2} \int_{S^{m-k-1}} f(x, \sqrt{1 - |x|^2} \theta) d\sigma_{m-k-1}(\theta) dx.$$

Let  $E = \text{span}\{e_1, \dots, e_k\}$ . Let  $H \subseteq \mathbb{R}^n$  be a random subspace of dimension  $n/2$  distributed uniformly. Let  $\lambda_1, \dots, \lambda_k$  be the singular values of the projection map  $P : H \rightarrow E$ . The singular values  $\lambda_1, \dots, \lambda_k$

are the cosines of the principal angles between  $H$  and  $E$ . The next proposition along with Proposition 3.1, allows us to study the distribution of singular values of a random matrix instead of integration on a random subspace.

**Proposition 3.2.** *Let  $H$  be a random subspace of dimension  $n/2$ , let  $E$  and  $\lambda_1, \dots, \lambda_k$  be as before. Let  $\Lambda = \text{diag}(\lambda_1, \dots, \lambda_k)$ . Let  $U : E \rightarrow E$  be a random orthogonal map distributed uniformly, independent of  $H$ . Let  $f : S^{n-1} \rightarrow \mathbb{R}$  be a measurable function such that  $f$  depends only on the first  $k$  coordinates. Then, the random variable*

$$\sqrt{\int_{S_H} f(x) d\sigma_H(x) \int_{S_{H^\perp}} f(x) d\sigma_{H^\perp}(x)}$$

is equal in distribution to

$$C_{n/2,k} \sqrt{\int_{B_k} f(U\Lambda U^T x) (1 - |x|^2)^{(n/2-k-2)/2} dx} \sqrt{\int_{B_k} f\left(U\sqrt{I - \Lambda^2} U^T x\right) (1 - |x|^2)^{(n/2-k-2)/2} dx}. \quad (2)$$

The constant  $C_{n/2,k}$  is the same as in Proposition 3.1 with  $m = n/2$ .

*Proof.* In this proof, we construct an orthogonal map  $V : \mathbb{R}^n \rightarrow \mathbb{R}^n$  that maps  $H$  and  $H^\perp$  to canonical subspaces that depend only on the principle angles and a rotation  $U : E \rightarrow E$ . The map  $V$  is chosen such that  $f$  would be invariant under  $V$ . In order to construct  $V$  we use the projection maps to define appropriate orthonormal bases for  $H$ ,  $H^\perp$  and  $E$ .

By the Singular Value Decomposition (SVD) of the projection  $P : H \rightarrow E$  there exists an orthonormal basis  $x_1, \dots, x_k$  of  $E$  and an orthonormal basis  $y_1, \dots, y_{n/2}$  of  $H$  such that

$$P = \sum_{i=1}^k \lambda_i x_i \otimes y_i.$$

Since  $P y_j = 0$  for all  $j = k+1, \dots, n/2$  we have  $y_{k+1}, \dots, y_{n/2} \in E^\perp$ . Since  $P y_i = \lambda_i x_i$  for  $i = 1, \dots, k$  there exists a unit vector  $v_{n/2+i} \in E^\perp$  such that

$$y_i = \lambda_i x_i + \sqrt{1 - \lambda_i^2} v_{n/2+i}, \quad \forall i = 1, \dots, k.$$

Denote  $v_j = y_j$  for  $j = k+1, \dots, n/2$ . Let  $i' = n/2 + i$  where  $i \leq k$  and let  $k+1 \leq j \leq n/2$ . Since  $v_{k+1}, \dots, v_{n/2+k} \in E^\perp$ , we have

$$\begin{aligned} 0 &= \langle y_i, y_j \rangle = \left\langle \lambda_i x_i + \sqrt{1 - \lambda_i^2} v_{i'}, v_j \right\rangle \\ &= \sqrt{1 - \lambda_i^2} \langle v_{i'}, v_j \rangle. \end{aligned}$$

With probability 1 we have  $0 < \lambda_i < 1$ , hence with probability 1

$$\langle v_{i'}, v_j \rangle = 0.$$

By the same argument we obtain  $\langle v_i, v_j \rangle = 0$  for all  $i \neq j$ . Let  $P_\perp : H^\perp \rightarrow E$  be the orthogonal projection to  $E$ . The singular values of  $P_\perp$  are exactly  $\sqrt{1 - \lambda_1^2}, \dots, \sqrt{1 - \lambda_k^2}$ . Note that

$$\sqrt{1 - \lambda_1^2} x_1 - \lambda_1 v_{n/2+1}, \dots, \sqrt{1 - \lambda_k^2} x_k - \lambda_k v_{n/2+k} \in H^\perp$$

are orthogonal to each other. Since the SVD is unique, up to trivial transformations, we have

$$P_\perp = \sum_{i=1}^k \sqrt{1 - \lambda_i^2} x_i \otimes \left( \sqrt{1 - \lambda_i^2} x_i - \lambda_i v_{n/2+i} \right).$$

Hence, there exist  $v_{n/2+k+1}, \dots, v_n \in E^\perp$  such that

$$\sqrt{1 - \lambda_1^2}x_1 - \lambda_1 v_{n/2+1}, \dots, \sqrt{1 - \lambda_k^2}x_k - \lambda_k v_{n/2+k}, v_{n/2+k+1}, \dots, v_n$$

is an orthonormal basis of  $H^\perp$ . By the same argument as before, with probability one,  $v_{n/2+1}, \dots, v_n$  are orthogonal to each other. Since  $v_{n/2+k+1}, \dots, v_n \in H^\perp$  and  $v_{k+1}, \dots, v_{n/2} \in H$  we find that  $x_1, \dots, x_k, v_{k+1}, \dots, v_n$  is an orthonormal basis of  $\mathbb{R}^n$ . Let  $V$  be the orthogonal map defined by  $Vx_i = x_i$  for  $i = 1, \dots, k$  and  $Vv_j = e_j$  for  $j = k+1, \dots, n$ . Since  $V$  is the identity map on  $E$  we have  $f(Vx) = f(x)$  for all  $x \in S^{n-1}$ . Hence,

$$\begin{aligned} \sqrt{\int_{S_H} f(x) d\sigma_H(x) \int_{S_{H^\perp}} f(x) d\sigma_{H^\perp}(x)} &= \sqrt{\int_{S_H} f(Vx) d\sigma_H(x) \int_{S_{H^\perp}} f(Vx) d\sigma_{H^\perp}(x)} \\ &= \sqrt{\int_{S_{VH}} f(x) d\sigma_{VH}(x) \int_{S_{VH^\perp}} f(x) d\sigma_{VH^\perp}(x)} \end{aligned}$$

Let  $B_{H,k}$  be the unit ball of

$$V(H \cap (E^\perp \cap H)^\perp) = \text{span}\{\lambda_1 x_1 + \sqrt{1 - \lambda_1^2} e_{n/2+1}, \dots, \lambda_k x_k + \sqrt{1 - \lambda_k^2} e_{n/2+k}\}.$$

For any

$$x = \sum_{i=1}^k t_i \left( \lambda_i x_i + \sqrt{1 - \lambda_i^2} e_{n/2+i} \right) \in B_{H,k},$$

where  $\sum_{i=1}^k t_i^2 \leq 1$ , we have

$$f(x) = f\left(\sum_{i=1}^k t_i \left( \lambda_i x_i + \sqrt{1 - \lambda_i^2} e_{n/2+i} \right)\right) = f\left(\sum_{i=1}^k t_i \lambda_i x_i\right).$$

Let  $U : E \rightarrow E$  be the orthogonal map defined by  $Ux_i = e_i$  for  $i = 1, \dots, k$ . We have,

$$\int_{B_{H,k}} f(x) dx = \int_{B_k} f(U\Lambda U^T x) dx,$$

where  $B_k$  is the unit ball of  $E$ . Since the distribution of  $H$  is invariant under the action of  $O(k) \times O(n-k)$ , the distribution of  $U$  is uniform over the orthogonal maps of  $E$ . Let  $B_{H^\perp,k}$  be the unit ball of  $\text{span}\{\sqrt{1 - \lambda_1^2}x_1 - \lambda_1 e_{n/2+1}, \dots, \sqrt{1 - \lambda_k^2}x_k - \lambda_k e_{n/2+k}\}$ . Using the same map  $U$ , we have

$$\int_{B_{H^\perp,k}} f(x) dx = \int_{B_k} f(U\sqrt{I - \Lambda^2} U^T x) dx.$$

To finish the proof we use Proposition 3.1 on  $S_{VH}$  and  $S_{VH^\perp}$ . □

With probability 1, the matrices  $\Lambda$  and  $\sqrt{I - \Lambda^2}$  are invertible. Hence, using the change of variables formula, (2) can be written as

$$\begin{aligned} \frac{C_{n/2,k}}{\sqrt{\prod_{j=1}^k \lambda_j \sqrt{1 - \lambda_j^2}}} &\sqrt{\int_{\mathbb{R}^k} f(x) \left(1 - |\Lambda^{-1} U^T x|^2\right)_+^{(n/2-k-2)/2} dx} \\ &\times \sqrt{\int_{\mathbb{R}^k} f(x) \left(1 - |(I - \Lambda^2)^{-1/2} U^T x|^2\right)_+^{(n/2-k-2)/2} dx}. \end{aligned}$$

By the Cauchy-Schwarz inequality this is at least

$$\frac{C_{n/2,k}}{\sqrt{\prod_{j=1}^k \lambda_j \sqrt{1-\lambda_j^2}}} \int_{\mathbb{R}^k} f(x) \left(1 - |\Lambda^{-1} U^T x|^2\right)_+^{(n/2-k-2)/4} \left(1 - |(I - \Lambda^2)^{-1/2} U^T x|^2\right)_+^{(n/2-k-2)/4} dx. \quad (3)$$

The random variables  $\lambda_1, \dots, \lambda_k$  are the singular values of a block of size  $n/2 \times k$  in a random orthogonal matrix. These singular values can be described using Wishart matrices [2]

**Proposition 3.3.** *Let  $N_1, N_2$  be  $(n/2) \times k$  independent random matrices with independent standard Gaussian entries. Let  $X$  be a random orthogonal matrix, chosen by the Haar uniform distribution. Let*

$$X = \begin{pmatrix} X_{1,1} & X_{1,2} \\ X_{2,1} & X_{2,2} \end{pmatrix}$$

Where  $X_{1,1}$  is  $(n/2) \times k$  block. Then, the singular values of  $X_{1,1}$  have the same distribution as the square roots of the eigenvalues of  $N_1^T N_1 (N_1^T N_1 + N_2^T N_2)^{-1}$ .

Upper and lower bounds for the eigenvalues of the above matrix, can be achieved using a concentration result by Gordon for singular values of Gaussian matrices [10].

**Lemma 3.4.** *Let  $A$  be  $(n/2) \times k$  random matrix with independent standard Gaussian entries. Assume that  $k \leq n/2$ . Let  $s_1 \leq \dots \leq s_k$  be the singular values of  $A$ , then with probability greater than  $1 - 2e^{-t^2/2}$  we have*

$$\sqrt{n/2} - \sqrt{k} - t \leq s_1 \leq s_k \leq \sqrt{n/2} + \sqrt{k} + t.$$

Combining both results, we have

**Proposition 3.5.** *Let  $k \leq \alpha_1 \sqrt{n}$  and let  $\lambda_1, \dots, \lambda_k$  be as before. Then, with probability greater than  $1 - 4e^{-\sqrt{n}}$ , we have*

$$\left| \lambda_i - \frac{1}{\sqrt{2}} \right| \leq \frac{C(\sqrt{\alpha_1} + \sqrt{2})}{n^{1/4}}, \quad \forall i = 1, \dots, k.$$

*Proof.* Let  $N_1, N_2$  be as in Proposition 3.3. By Lemma 3.4 there exists  $\mu_1, \dots, \mu_k, \sigma_1, \dots, \sigma_k$  and  $U, V$  orthogonal matrices such that

$$N_1^T N_1 = U \text{diag}(\mu_1^2, \dots, \mu_k^2) U^T, \quad N_2^T N_2 = V \text{diag}(\sigma_1^2, \dots, \sigma_k^2) V^T,$$

and, there exists  $C' > 0$  such that, with probability greater than  $1 - 4e^{-\sqrt{n}}$ ,

$$\left| \mu_i^2 - \frac{n}{2} \right|, \left| \sigma_i^2 - \frac{n}{2} \right| \leq C'(\sqrt{\alpha_1} + \sqrt{2})n^{3/4}, \quad (4)$$

for all  $i = 1, \dots, k$ . Assume that event (4) holds true. Let  $E_1, E_2$  be defined by  $N_1^T N_1 = (n/2)I + E_1$  and  $N_2^T N_2 = (n/2)I + E_2$ . Then  $\|E_i\|_{op} \leq C_i(\sqrt{\alpha_1} + \sqrt{2})n^{3/4}$  for  $i = 1, 2$ . We have,

$$N_1^T N_1 (N_1^T N_1 + N_2^T N_2)^{-1} = \frac{1}{2} \left( I + \frac{2}{n} E_1 \right) \left( I + \frac{1}{n} (E_1 + E_2) \right)^{-1}.$$

Let  $T_1 = (E_1 + E_2)/n$  and  $T_2 = E_1/n$ , then  $\|T_i\|_{op} \leq C'_i(\sqrt{\alpha_1} + \sqrt{2})n^{-1/4}$  for  $i = 1, 2$ . We have

$$N_1^T N_1 (N_1^T N_1 + N_2^T N_2)^{-1} = \left( \frac{1}{2} I + T_2 \right) \left( I - T_1 + \sum_{j=2}^{\infty} (-1)^j T_1^j \right).$$

Hence,

$$N_1^T N_1 (N_1^T N_1 + N_2^T N_2)^{-1} = \frac{1}{2} I + T,$$

where  $\|T\|_{op} \leq \tilde{C}(\sqrt{\alpha_1} + \sqrt{2})n^{-1/4}$ . □



**Corollary 3.6.** *With probability greater than  $1 - 4e^{-\sqrt{n}}$  we have*

$$\left| \left| U\Lambda^{-1}U^T x \right|^2 + \left| U(I - \Lambda^2)^{-1/2}U^T x \right|^2 - 4|x|^2 \right| \leq C(\sqrt{\alpha_1} + \sqrt{2})^2 |x|^2 / \sqrt{n}, \quad \forall x \in \mathbb{R}^k.$$

*Proof.* Assume that

$$\Lambda = \frac{1}{\sqrt{2}}I + T,$$

where  $\|T\|_{op} \leq C'(\sqrt{\alpha_1} + \sqrt{2})/n^{1/4}$ . By the above proposition, this event has probability greater than  $1 - 4e^{-\sqrt{n}}$ . We have,

$$\begin{aligned} \Lambda^{-2} + (I - \Lambda^2)^{-1} &= \left( \frac{1}{2}I + \sqrt{2}T + T^2 \right)^{-1} + \left( \frac{1}{2}I - \sqrt{2}T - T^2 \right)^{-1} \\ &= 4(I - 4(\sqrt{2}T + T^2)^2)^{-1} = 4I + \tilde{T}, \end{aligned}$$

where  $\|\tilde{T}\|_{op} \leq \tilde{C}\|T^2\|_{op} \leq C(\sqrt{\alpha_1} + \sqrt{2})^2/\sqrt{n}$ . Hence,

$$\begin{aligned} \left| U\Lambda^{-1}U^T x \right|^2 + \left| U(I - \Lambda^2)^{-1/2}U^T x \right|^2 &= \langle U(\Lambda^{-2} + (I - \Lambda^2)^{-1})U^T x, x \rangle \\ &= 4|x|^2 + \langle U\tilde{T}U^T x, x \rangle. \end{aligned}$$

Hence,

$$\left| \left| U\Lambda^{-1}U^T x \right|^2 + \left| U(I - \Lambda^2)^{-1/2}U^T x \right|^2 - 4|x|^2 \right| \leq \|\tilde{T}\|_{op} |x|^2 \leq C(\sqrt{\alpha_1} + \sqrt{2})^2 |x|^2 / \sqrt{n}.$$

□

The above proof demonstrates how considering both  $H$  and  $H^\perp$  simultaneously can cancel the first order term in concentration inequalities. This cancellation leads to great improvement of the estimations, and it is one of the fundamental ideas of our approach.

The concentration of the principal angles, allows us to evaluate the coefficient  $C_{n/2,k} \left( \prod_{j=1}^k \lambda_j \sqrt{1 - \lambda_j^2} \right)^{-1/2}$  and the integral at (3).

**Proposition 3.7.** *Let  $C_{n,k}$  and  $C_{n/2,k}$  be the constants from Proposition 3.1 with  $m = n, n/2$ . For  $k \leq \alpha_1 \sqrt{n}$ , we have*

$$2^{k/2} \frac{C_{n/2,k}}{C_{n,k}} \geq Ce^{-\alpha_1^2/4}.$$

*Proof.* By the definition of  $C_{n,k}$  and  $C_{n/2,k}$ , we need to estimate

$$2^{k/2} \frac{\Gamma(n/4 + 1/2)\Gamma(n/2 - k/2 + 1/2)}{\Gamma(n/4 - k/2 + 1/2)\Gamma(n/2 + 1/2)}$$

Using Sterlings formula and the assumption on  $k$ , this is

$$\left( 1 + O\left( \frac{1}{\sqrt{n}} \right) \right) \exp\left( -\frac{k^2}{4n} + O\left( \frac{1}{\sqrt{n}} \right) \right).$$

□

Hence, by choosing  $\alpha_1$  small enough and using the concentration result for  $\lambda_i \sqrt{1 - \lambda_i^2}$  (as in Corollary 3.6) we have:

**Corollary 3.8.** *Let  $k \leq \alpha_1 \sqrt{n}$  and let  $\lambda_1, \dots, \lambda_k$  be as before. Then, with probability greater than  $1 - 4e^{-\sqrt{n}}$ , we have*

$$C_{n/2,k} \sqrt{\prod_{j=1}^k \frac{1}{\lambda_j \sqrt{1 - \lambda_j^2}}} \geq 0.98 C_{n,k}.$$

*Proof.* Assume that for all  $1 \leq i \leq k$  we have  $\lambda_i^2 = 1/2 + t_i$  where  $|t_i| \leq C'(\sqrt{\alpha_1} + \sqrt{2})/n^{1/4}$ . By Proposition 3.5, this event has probability greater than  $1 - 4e^{-\sqrt{n}}$ . We have,

$$\frac{1}{\lambda_i \sqrt{1 - \lambda_i^2}} = \sqrt{\frac{1}{(1/2 + t_i)(1/2 - t_i)}} = \frac{2}{\sqrt{1 - 4t_i^2}}.$$

Hence,

$$\left| \frac{1}{\lambda_i \sqrt{1 - \lambda_i^2}} - 2 \right| \leq \frac{C(\sqrt{\alpha_1} + \sqrt{2})^2}{\sqrt{n}}.$$

We have

$$\sqrt{\prod_{j=1}^k \frac{1}{\lambda_j \sqrt{1 - \lambda_j^2}}} \geq 2^{k/2} \exp \left( -\frac{C(\sqrt{\alpha_1} + \sqrt{2})^2 k}{4\sqrt{n}} + O \left( \frac{k(\sqrt{\alpha_1} + \sqrt{2})^2}{n} \right) \right).$$

We may assume that  $\alpha_1$  is small enough, such that both

$$\exp \left( -\frac{C(\sqrt{\alpha_1} + \sqrt{2})^2 \alpha_1}{4} + O \left( \frac{\alpha_1(\sqrt{\alpha_1} + \sqrt{2})^2}{\sqrt{n}} \right) \right) \geq 0.99,$$

and (By Proposition 3.7),  $0.99 C_{n/2,k} 2^{k/2} \geq 0.98 C_{n,k}$ . Hence,

$$C_{n/2,k} \sqrt{\prod_{j=1}^k \frac{1}{\lambda_j \sqrt{1 - \lambda_j^2}}} \geq 0.99 C_{n/2,k} 2^{k/2} \geq 0.98 C_{n,k}.$$

□

In order to understand the integral in (3), we write  $\mathbb{R}^k$  as  $\rho n^{-1/4} B_k \cup (\mathbb{R}^k \setminus \rho n^{-1/4} B_k)$  where  $\rho > 0$ . Inside the ball  $\rho n^{-1/4} B_k$  the integral is close to 1. Outside the ball, we show that the integral is negligible.

The estimation inside the ball of radius  $\rho n^{-1/4}$  uses standard inequalities and corollary 3.6 (see Appendix B for the proof). In this proposition we define an upper bound on  $\rho$ .

**Proposition 3.9.** *Let  $k \leq \alpha_1 \sqrt{n}$ . Let  $\Lambda$  and  $U$  be as before. Then with probability greater than  $1 - 4e^{-\sqrt{n}}$*

$$\begin{aligned} & \int_{\rho n^{-1/4} B_k} f(x) \left( 1 - |\Lambda^{-1} U^T x|^2 \right)_+^{(n/2-k-2)/4} \left( 1 - |(I - \Lambda^2)^{-1/2} U^T x|^2 \right)_+^{(n/2-k-2)/4} dx \\ & \geq 0.95 \int_{\rho n^{-1/4} B_k} f(x) \left( 1 - |x|^2 \right)_+^{(n-k-2)/2} dx. \end{aligned}$$

Using the Laplace method (see Appendix B), we estimate the integral outside  $\rho n^{-1/4} B_k$ .

**Proposition 3.10.** *Let  $k \leq \alpha_1 \sqrt{n}$ . Then, for any  $f : \mathbb{R}^n \rightarrow \mathbb{R}_+$  with  $\|f\|_\infty \leq e^{\alpha_2 \sqrt{n}}$ , we have*

$$I = C_{n,k} \int_{\mathbb{R}^k \setminus \rho n^{-1/4} B_k} f(x) (1 - |x|^2)_+^{(n-k-2)/2} dx \leq \frac{2\alpha_1}{\rho^2} e^{-\alpha_2 \sqrt{n}}.$$

The constant  $C_{n,k}$  is the same as in Proposition 3.1, and  $\rho > 0$  is the same as in Proposition 3.9.

*Proof of Theorem 1.2.* By Proposition 3.2 and the Cauchy-Schwarz inequality, the event

$$\sqrt{\int_{S_H} f(x) d\sigma_H(x) \int_{S_{H^\perp}} f(x) d\sigma_{H^\perp}(x)} \geq 0.9$$

has the greater probability than the event

$$\frac{C_{n/2,k}}{\sqrt{\prod_{j=1}^k \lambda_j \sqrt{1 - \lambda_j^2}}} \int_{\mathbb{R}^k} f(x) (1 - |\Lambda^{-1} U^T x|^2)_+^{(n/2-k-2)/4} \left(1 - \left|(I - \Lambda^2)^{-1/2} U^T x\right|^2\right)_+^{(n/2-k-2)/4} dx \geq 0.9.$$

By Corollary 3.8 and Proposition 3.9 with probability greater than  $1 - 4e^{-\sqrt{n}}$  the left hand side is at least

$$0.93 C_{n,k} \int_{\rho n^{-1/4} B_k} f(x) (1 - |x|^2)_+^{(n-k-2)/2} dx.$$

By Proposition 3.1 this is equal to

$$0.93 \left( \int_{S^{n-1}} f(x) d\sigma_{n-1}(x) - C_{n,k} \int_{\mathbb{R}^k \setminus \rho n^{-1/4} B_k} f(x) (1 - |x|^2)_+^{(n-k-2)/2} dx \right).$$

By Proposition 3.10 there exists  $\hat{C} > 0$  such that for all  $n > \hat{C}$  we have

$$0.93 C_{n,k} \int_{\mathbb{R}^k \setminus \rho n^{-1/4} B_k} f(x) (1 - |x|^2)_+^{(n-k-2)/2} dx \leq 0.01.$$

Hence,

$$0.93 \left( \int_{S^{n-1}} f(x) d\sigma_{n-1}(x) - C_{n,k} \int_{\mathbb{R}^k \setminus \rho n^{-1/4} B_k} f(x) (1 - |x|^2)_+^{(n-k-2)/2} dx \right) \geq 0.9.$$

□

## A Protocol for VSP

The protocol we present here is a simple modification of the one presented by Raz [8].

As before,  $c, c_1, C$  etc. denote positive universal constants.

Let  $k = \lfloor e^{\sqrt{n}} \rfloor$ . Let  $E_1, \dots, E_k \subseteq \mathbb{R}^n$  be independent random subspaces of dimension  $\lfloor C_1 \sqrt{n} \rfloor$  chosen uniformly. For every  $1 \leq i \leq k$ , let  $\mathcal{N}_i = \{\theta_1^i, \dots, \theta_m^i\}$  be independent random vectors in  $S^{n-1} \cap E_i$ , where  $m = \lfloor e^{C_2 \sqrt{n}} \rfloor$ . Alice and Bob sample  $(E_1, \mathcal{N}_1), \dots, (E_k, \mathcal{N}_k)$  in advance and store the results. Each real number stored by Alice and Bob is kept with accuracy of  $\log n$  bits. The protocol will be the following: Alice chooses a random index  $1 \leq \hat{i} \leq k$ , and then finds the index  $1 \leq \hat{j} \leq m$  such that

$$\max_{1 \leq j \leq m} \langle \theta_j^{\hat{i}}, u \rangle = \langle \theta_{\hat{j}}^{\hat{i}}, u \rangle.$$

Alice sends Bob both indices  $\hat{i}$  and  $\hat{j}$  using at most  $\log k + \log m = (1 + C_2)\sqrt{n}$  bits. Bob checks the distance of  $\theta_{\hat{j}}^i$  to  $H$  and  $H^\perp$ . If  $d(\theta_{\hat{j}}^i, H) > d(\theta_{\hat{j}}^i, H^\perp)$  then they answer that  $u \in H$  otherwise they answer that  $u \in H^\perp$ .

In this protocol Alice performs one measurement. This measurement is in a subspace of dimension  $O(\sqrt{n})$ , hence the protocol has total rank of  $O(\sqrt{n})$ .

The analysis of this protocol is done in two steps. First we show that the protocol works when we replace  $(E_1, \mathcal{N}_1), \dots, (E_k, \mathcal{N}_k)$  with shared random pair  $(E, \mathcal{N})$ . The complexity of the public coin protocol is  $\log m = C_2\sqrt{n}$  bits. Second, we eliminate the need for shared randomness by considering  $(E_1, \mathcal{N}_1), \dots, (E_k, \mathcal{N}_k)$ . This step is standard, and the cost of eliminating the shared randomness is another  $\log k = \sqrt{n}$  bits. We present the main ideas of these steps.

In the analysis of the first step, we use two standard results: In the first, we use the fact that the norm of a projection of a random vector is close to Gaussian [3].

**Proposition A.1.** *Let  $v \in S^{d-1}$  be a random vector distributed uniformly. Let  $F \subseteq \mathbb{R}^d$  be a subspace of dimension  $\ell$ . Then,*

$$\mathbb{P}\left(\left|\text{Proj}_F v\right|^2 - \frac{\ell}{d}\right| \geq t\right) \leq C e^{-ct^2 d}, \quad \forall t.$$

Note that by applying a random rotation we may assume that  $v$  is fixed and  $F$  is random. The second standard result shows that our choice of  $\mathcal{N}_i$  is typically an  $1/2$ -net of  $S^{n-1} \cap E_i$ .

**Proposition A.2.** *Let  $z_1, \dots, z_\ell$  be independent uniformly chosen random vectors in  $S^{d-1}$ , where  $\ell = e^{Cd}$ . Then with probability greater than  $1 - e^{-e^{c\ell}}$  they form an  $1/2$ -net of the sphere.*

*Sketch of the proof.* Let  $N$  be an  $\varepsilon$ -net with  $\#N = e^{c_1 k}$  (e.g [7]). For any  $x \in N$  we have

$$\mathbb{P}(|z_i - x| > \varepsilon \forall i) \leq e^{-m} \mathbb{P}(|z_1 - x| \leq \varepsilon).$$

Since

$$\mathbb{P}(|z_1 - x| \leq \varepsilon) \approx e^{-c_2(1-\varepsilon)^2 k},$$

we have

$$\mathbb{P}(\exists x \in N; |z_i - x| > \varepsilon \forall i) \leq \exp\left(c_1 k - e^{(c - c_2(1-\varepsilon)^2)k}\right).$$

□

We are now ready to prove that the protocol works with a shared random pair  $(E, \mathcal{N})$ .

*Proof.* Let  $u, H$  be fixed, such that either  $u \in H$  or  $u \in H^\perp$ . We have,

$$\max_{\theta \in S^{n-1} \cap E} \langle u, \theta \rangle = \max_{\theta \in S^{n-1} \cap E} \langle u, \text{Proj}_E \theta \rangle = \max_{\theta \in S^{n-1} \cap E} \langle \text{Proj}_E u, \theta \rangle = |\text{Proj}_E u|.$$

The dimension of  $E$  is  $\lfloor C_1 \sqrt{n} \rfloor$ . By Proposition A.1 and (??) we can choose  $C_1$  big enough such that

$$\mathbb{P}\left(\max_{\theta \in S^{n-1} \cap E} \langle u, \theta \rangle \geq \frac{100}{n^{1/4}}\right) \geq 0.91.$$

Let  $\theta_j \in \mathcal{N}$  be the closest point to  $u$ . By Proposition A.2, with probability greater than 0.99, the set  $\mathcal{N} = \{\theta_1, \dots, \theta_m\}$  is an  $1/2$ -net of  $S^{n-1} \cap E$ . Hence, for any  $\theta \in S^{n-1} \cap E$  there exists  $\theta_i \in \mathcal{N}$  such that  $|\theta_i - \theta| \leq 1/2$ . Hence,

$$\langle \theta, u \rangle = \langle \theta - \theta_i, u \rangle + \langle \theta_i, u \rangle \leq |\theta - \theta_i| |\text{Proj}_E u| + \max_j \langle \theta_j, u \rangle \leq \frac{1}{2} |\text{Proj}_E u| + \max_j \langle \theta_j, u \rangle.$$

The right hand side does not depend on  $\theta$ , hence,

$$\max_j \langle \theta_j, u \rangle \geq \max_{\theta \in S^{n-1} \cap E} \langle \theta, u \rangle - \frac{1}{2} |\text{Proj}_E u| = \frac{1}{2} \max_{\theta \in S^{n-1} \cap E} \langle \theta, u \rangle.$$

Let  $\alpha = \langle \theta_{\hat{j}}, u \rangle$ . With probability greater than 0.9 we have

$$\alpha \geq \frac{50}{n^{1/4}}.$$

Let

$$\theta_{\hat{j}} = \alpha u + \sqrt{1 - \alpha^2} v,$$

where  $v \in S^{n-1} \cap u^\perp$ . By the definition  $v$ , it is distributed uniformly in  $S^{n-1} \cap u^\perp$ . By Proposition A.1 we have

$$\mathbb{P} \left( \left| |\text{Proj}_H v|^2 - \frac{1}{2} \right| \geq \frac{10}{\sqrt{n}} \right) \leq 0.1.$$

Hence, if  $u \in H$ , then with probability greater than 0.8 we have,

$$|\text{Proj}_H \theta_{\hat{j}}|^2 = \alpha^2 + (1 - \alpha^2) |\text{Proj}_H v|^2 \geq \frac{1}{2} + \frac{1000}{\sqrt{n}},$$

and

$$|\text{Proj}_{H^\perp} \theta_{\hat{j}}|^2 \leq |\text{Proj}_{H^\perp} v|^2 \leq \frac{1}{2} + \frac{10}{\sqrt{n}}.$$

Hence, with probability greater than 0.8 we have  $|\text{Proj}_{H^\perp} \theta_{\hat{j}}| < |\text{Proj}_H \theta_{\hat{j}}|$ , thus the protocol would correctly determine that  $u \in H$ . The case  $u \in H^\perp$  is proven similarly.  $\square$

Next we explain how to eliminate the shared randomness.

*Proof.* We denote by  $\theta_j \in \mathcal{N}$  the closest vector to  $u$  in  $\mathcal{N}$ . Let

$$A = \left\{ (u, H, E, \mathcal{N}); \begin{array}{l} |\text{Proj}_H \theta_j|^2 > |\text{Proj}_{H^\perp} \theta_j|^2 + 10/\sqrt{n}, \text{ if } u \in H \\ |\text{Proj}_{H^\perp} \theta_j|^2 > |\text{Proj}_H \theta_j|^2 + 10/\sqrt{n}, \text{ if } u \in H^\perp \end{array} \right\}.$$

Let  $A_{u,H}$  and  $A_{E,\mathcal{N}}$  denote the corresponding sections of the set  $A$ . By the previous step of shared randomness, for any fixed  $(u, H)$  we have,

$$\mathbb{P}_{(E,\mathcal{N})} ((u, H) \in A_{E,\mathcal{N}}) \geq 0.8.$$

By the Chernoff-Hoeffding inequality, for any fixed  $u, H$  we have

$$\mathbb{P} \left( \frac{\#\{i; (E_i, \mathcal{N}_i) \in A_{u,H}\}}{m} \leq 0.8 \right) \leq e^{-ck}.$$

Hence, by Fubini's theorem, for most choices of  $(E_1, \mathcal{N}_1), \dots, (E_k, \mathcal{N}_k)$

$$\mathbb{P}_{u,H} \left( \frac{\#\{i; (E_i, \mathcal{N}_i) \in A_{u,H}\}}{m} \leq 0.8 \right) \leq e^{-c'k}.$$

Recall that Alice and Bob sample in advance the list  $(E_1, \mathcal{N}_1), \dots, (E_k, \mathcal{N}_k)$ . Thus, with high probability, their protocol works for a any  $(u, H)$  outside a set of measure  $e^{-c'e\sqrt{n}}$ . Hence, for any vector  $u$  and a subspace  $H$  we can find  $u'$  and  $H'$  for which the protocol works,  $|u - u'| \leq 1/\sqrt{n}$  and  $\|\text{Proj}_H - \text{Proj}_{H'}\|_{op} \leq 1/\sqrt{n}$ . Therefore  $|\text{Proj}_H u - \text{Proj}_{H'} u'| \leq 2/\sqrt{n}$  and the protocol works for arbitrary  $u$  and  $H$ .  $\square$

## B Asymptotic estimates

Here we present the proofs of Propositions 3.9 and 3.10.

*proof of Proposition 3.9.* Assume the event  $\|\Lambda - I/\sqrt{2}\|_{op} \leq C(\sqrt{\alpha_1} + \sqrt{2})/n^{1/4}$  holds true. By Proposition 3.5 this event has probability greater than  $1 - 4e^{-\sqrt{n}}$ . Define  $\psi : \mathbb{R}^n \rightarrow \mathbb{R}$  by

$$\psi(x) = \left(1 - |\Lambda^{-1}U^T x|^2\right)_+^{(n/2-k-2)/4} \left(1 - \left|(I - \Lambda^2)^{-1/2}U^T x\right|^2\right)_+^{(n/2-k-2)/4}.$$

Using the Taylor expansion

$$\log(1 - |x|^2) = -|x|^2 + O(|x|^4)$$

for any  $|x| < 3/4$ , we have

$$\psi(x) = \exp\left(-(n/8 - k/4 - 1/2)\left(|U\Lambda^{-1}U^T x|^2 + |U(I - \Lambda^2)^{-1/2}U^T x|^2 + O(|x|^4)\right)\right),$$

for any  $|x| \leq 1/10$ . By Corollary 3.6 for any  $x \in \rho n^{-1/4}B_k$  we have

$$\psi(x) = \exp\left(-(n - k - 2)|x|^2/2 + O(\rho^2(\sqrt{\alpha_1} + \sqrt{2})^2) + O(\rho^4) + O((\alpha_1 + 1/\sqrt{n})\rho^2)\right).$$

In Corollary 3.8 we assumed an upper bound on  $\alpha_1$ . Under this upper bound assumption we can choose  $\rho > 0$  small enough, independent of  $n$  and any specific choice of  $\alpha_1$  such that

$$\psi(x) \geq 0.95 \exp(-(n/2 - k/2 - 1)|x|^2) \geq 0.95(1 - |x|^2)^{(n-k-2)/2}, \quad \forall x \in \rho n^{-1/4}B_k.$$

□

*proof of Proposition 3.10.* By the assumption on  $f$  we have

$$I \leq C_{n,k} e^{\alpha_2 \sqrt{n}} \int_{\mathbb{R}^k \setminus \rho n^{-1/4}B_k} (1 - |x|^2)_+^{(n-k-2)/2} dx$$

Using  $1 - x \leq e^{-x}$  and the assumption  $k \leq \alpha_1 \sqrt{n}$  and that  $\alpha_1$  is bounded by some universal constant, for  $n$  exceeding some universal constant, we have

$$I \leq C_{n,k} e^{\alpha_2 \sqrt{n}} \int_{\mathbb{R}^k \setminus \rho n^{-1/4}B_k} e^{-(n/2-k/2-1)|x|^2} dx \leq C_{n,k} e^{\alpha_2 \sqrt{n}} \int_{\mathbb{R}^k \setminus \rho n^{-1/4}B_k} e^{-n|x|^2/3} dx.$$

By integrating in polar coordinates, we have

$$I \leq C_{n,k} k \text{Vol}(B_k) e^{\alpha_2 \sqrt{n}} \int_{\rho n^{-1/4}}^{\infty} r^{k-1} e^{-nr^2/3} dr = C_{n,k} k \text{Vol}(B_k) e^{\alpha_2 \sqrt{n}} \frac{1}{n^{k/2}} \int_{\rho n^{1/4}}^{\infty} r^{k-1} e^{-r^2/3} dr.$$

Define  $h(r) = -(k-1) \log r + r^2/3$ . The function  $h$  is convex, hence

$$h(r) \geq h(\rho n^{1/4}) + h'(\rho n^{1/4})(r - \rho n^{1/4}).$$

Assuming  $\alpha_1 \leq \rho^2/6$  we have,

$$h'(\rho n^{1/4}) = -\frac{k-1}{\rho n^{1/4}} + \frac{2}{3}\rho n^{1/4} \geq \frac{1}{2}\rho n^{1/4}.$$

We have,

$$\int_{\rho n^{1/4}}^{\infty} e^{-h(r)} dr \leq e^{-h(\rho n^{1/4})} \int_{\rho n^{1/4}}^{\infty} e^{-\rho n^{1/4}(r-\rho n^{1/4})/2} dr = \frac{2}{\rho n^{1/4}} e^{-h(\rho n^{1/4})}.$$

Hence,

$$I \leq 2C_{n,k} \text{Vol}(B_k) \rho^{k-2} k n^{-k/4-1/2} e^{-\sqrt{n}(\rho^2/3-\alpha_2)}.$$

Using

$$C_{n,k} \text{Vol}(B_k) = \frac{n-k}{n} \frac{\text{Vol}(B_{n-k}) \text{Vol}(B_k)}{\text{Vol}(B_n)} = \frac{n-k}{n} \binom{n/2}{k/2} \leq \left(\frac{n \cdot e}{k}\right)^{k/2},$$

we have

$$I \leq \frac{2}{\rho^2} (\sqrt{e}\rho)^k \frac{n^{k/4-1/2}}{k^{k/2-1}} e^{-\sqrt{n}(\rho^2/3-\alpha_2)}.$$

Assuming  $\alpha_1 > 0$  is small enough such that  $\rho^2/3 - \alpha_1 \log(\rho\sqrt{e/\alpha_1}) > 0$ , we optimize over  $k$ , and get

$$I \leq \frac{2\alpha_1}{\rho^2} \exp \left[ -\sqrt{n} \left( \rho^2/3 - \alpha_1 \log(\rho\sqrt{e/\alpha_1}) - \alpha_2 \right) \right].$$

Hence, we can choose

$$2\alpha_2 < \rho^2/3 - \alpha_1 \log(\rho\sqrt{e/\alpha_1}),$$

to finish the proof. □

## References

- [1] Laszlo Babai, Peter Frankl, and Janos Simon. Complexity classes in communication complexity theory. In *Proceedings of the 27th Annual Symposium on Foundations of Computer Science, SFCS '86*, pages 337–347, Washington, DC, USA, 1986. IEEE Computer Society.
- [2] Alan Edelman and Brian D. Sutton. The beta-Jacobi matrix model, the CS decomposition, and generalized singular value problems. *Found. Comput. Math.*, 8(2):259–285, 2008.
- [3] William B. Johnson and Joram Lindenstrauss. Extensions of Lipschitz mappings into a Hilbert space. In *Conference in modern analysis and probability (New Haven, Conn., 1982)*, volume 26 of *Contemp. Math.*, pages 189–206. Amer. Math. Soc., Providence, RI, 1984.
- [4] Bo'az Klartag and Oded Regev. Quantum one-way communication can be exponentially stronger than classical communication. In *STOC'11—Proceedings of the 43rd ACM Symposium on Theory of Computing*, pages 31–40. ACM, New York, 2011.
- [5] Bo'az Klartag and Oded Regev. The vector in subspace problem. Slides for an Oberwolfach workshop, 2011.
- [6] Ilan Kremer. Master's thesis.
- [7] Gilles Pisier. *The Volume of Convex Bodies and Banach Space Geometry*. Cambridge University Press, Cambridge, 10 1989.
- [8] Ran Raz. Exponential separation of quantum and classical communication complexity. In *Annual ACM Symposium on Theory of Computing (Atlanta, GA, 1999)*, pages 358–367 (electronic). ACM, New York, 1999.

- [9] Alexander A. Razborov. Communication complexity. In *An Invitation to Mathematics*, pages 97–117. Springer-Verlag Berlin Heidelberg, 2011.
- [10] Roman Vershynin. Introduction to the non-asymptotic analysis of random matrices. In *Compressed sensing*, pages 210–268. Cambridge Univ. Press, Cambridge, 2012.
- [11] Andrew Chi Chih Yao. Probabilistic computations: toward a unified measure of complexity (extended abstract). In *18th Annual Symposium on Foundations of Computer Science (Providence, R.I., 1977)*, pages 222–227. IEEE Comput. Sci., Long Beach, Calif., 1977.
- [12] Andrew Chi-Chih Yao. Some complexity questions related to distributive computing (preliminary report). In *Proceedings of the Eleventh Annual ACM Symposium on Theory of Computing, STOC '79*, pages 209–213, New York, NY, USA, 1979. ACM.